SEQUENCE LISTING

5	(1) GENE	RAL INFORMATION:
,	(i)	APPLICANT: Chen, J. Don Li, Hui
10	(ii)	TITLE OF INVENTION: Transcriptional Coactivator for Nuclear Hormone Receptors
	(iii)	NUMBER OF SEQUENCES: 2
15	(iv)	CORRESPONDENCE ADDRESS: (A) ADDRESSEE: Lahive and Cockfield (B) STREET: 28 State Street (C) CITY: Boston (D) STATE: MA
20		(E) COUNTRY: USA (F) ZIP: 02109
□ □ □ □ □	(v)	COMPUTER READABLE FORM: (A) MEDIUM TYPE: Floppy disk (B) COMPUTER: IBM PC compatible (C) OPERATING SYSTEM: PC-DOS/MS-DOS (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
4 4 3 0	(vi)	CURRENT APPLICATION DATA: (A) APPLICATION NUMBER: (B) FILING DATE: (C) CLASSIFICATION:
	(viii)	ATTORNEY/AGENT INFORMATION: (A) NAME: Liepmann, W. Hugo (B) REGISTRATION NUMBER: 20,407 (C) REFERENCE/DOCKET NUMBER: UMM-026
	(ix)	TELECOMMUNICATION INFORMATION: (A) TELEPHONE: 617-227-7400
40		(B) TELEFAX: 617-742-4214
	(2) INFO	RMATION FOR SEQ ID NO:1:
45	(i)	SEQUENCE CHARACTERISTICS: (A) LENGTH: 4496 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear
50	(ii)	MOLECULE TYPE: cDNA
55	(ix)	FEATURE: (A) NAME/KEY: CDS (B) LOCATION: 864338

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

5	GCT	GAT	GT (GGACT	rcaga	AG AG	CCAA:	raaa?	A ATA	AAAC'	rgct	TGA	ACATO	CCT 7	rtgac	CTGGTT		60
	AGC	CAGT	rgc '	TGATO	TATE	AT TO	CAAG			GGA Gly							1	112
10			_	AGT Ser													1	L60
15				CTT Leu													2	208
20				ATT Ile 45													2	256
口 切 口 正25				AAT Asn													3	304
型 二 二 二 二 二 二 二 二 二 二 二 二 二 二 二 二 二 二 二				AGA Arg													3	352
- 30 □ □ □ □				GAT Asp													4	100
当 山35 山 の				GAT Asp													4	148
40				CTA Leu 125													4	196
45				GTC Val													5	544
				GTT Val													5	92
50				TTA Leu													6	540
55				AGA Arg													6	88

						ATT Ile											736
5						ACA Thr											784
10						GGG Gly											832
15						ACA Thr 255											880
20						CAT His									_		928
						TCC Ser											976
口 口 二 二 二 二 二 二 二 二 二 二 二 二 二 二 二 二 二 二				_	_	AGA Arg							_	_			1024
<u> </u>						TAT Tyr											1072
₩ ₩ ₩35						TTC Phe 335											1120
40	_	_				CTC Leu	_				_	_					1168
						CAC His											1216
45						GTT Val											1264
50						GGC Gly											1312
55						AGG Arg 415											1360
	GGG	CAG	ATG	AGT	GGA	GCT	AGG	TAT	GGĢ	GGT _.	TCC	AGT	AAC	ATA	GCT	TCA	1408

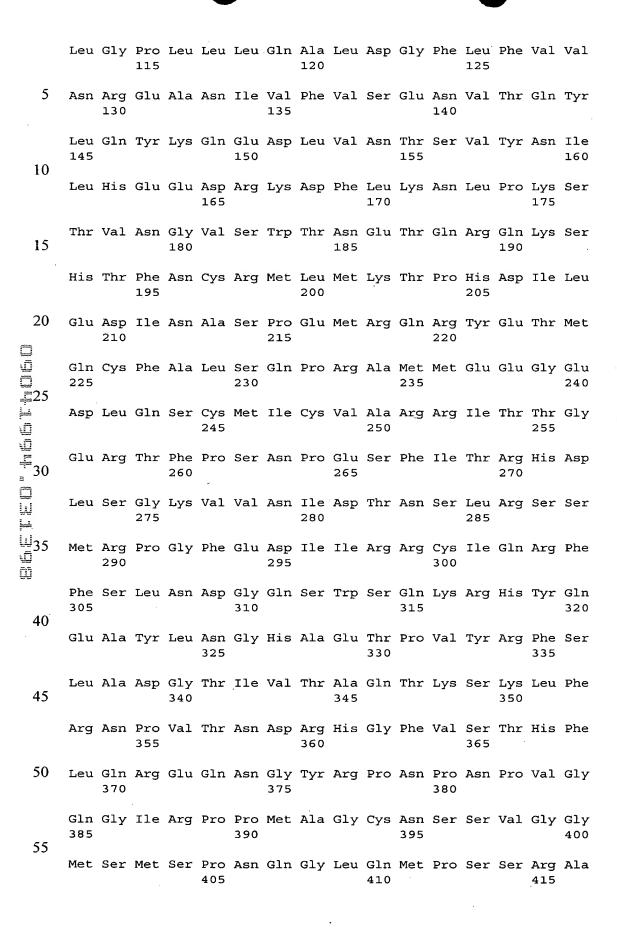
	Gly	Gln	Met	Ser	Gly 430	Ala	Arg	Tyr	Gly	Gly 435	Ser	Ser	Asn	Ile	Ala 440	Ser	
5					CCA Pro												1456
10					AAC Asn												1504
15					CAG Gln												1552
					TCA Ser												1600
20					TCT Ser 510												1648 ·
回 回 三 字 25 二 二 二 二 二 二 二 二 二 二 二 二 二 二 二 二 二 二					CTG Leu												1696
3					TCA Ser												1744
다 님 님 님35					CAA Gln												1792
					TAT Tyr												1840
40					AGA Arg 590												1888
45					GCA Ala												1936
50					CTG Leu												1984
55					ACC Thr												2032
					ACC Thr												2080

	650			655			660			665	
5			ACA Thr 670								2128
10			CAC His								2176
			ACT Thr		,						2224
15			GAC Asp								2272
20			AAT Asn								2320
			GCA Ala 750								2368
——————————————————————————————————————			ATG Met								2416
			GAC Asp								2464
山35 山 四			GAT Asp								2512
40			AAT Asn								2560
45			GTG Val 830								2608
50			CAG Gln								2656
			GTT Val								2704
55			ATG Met								2752

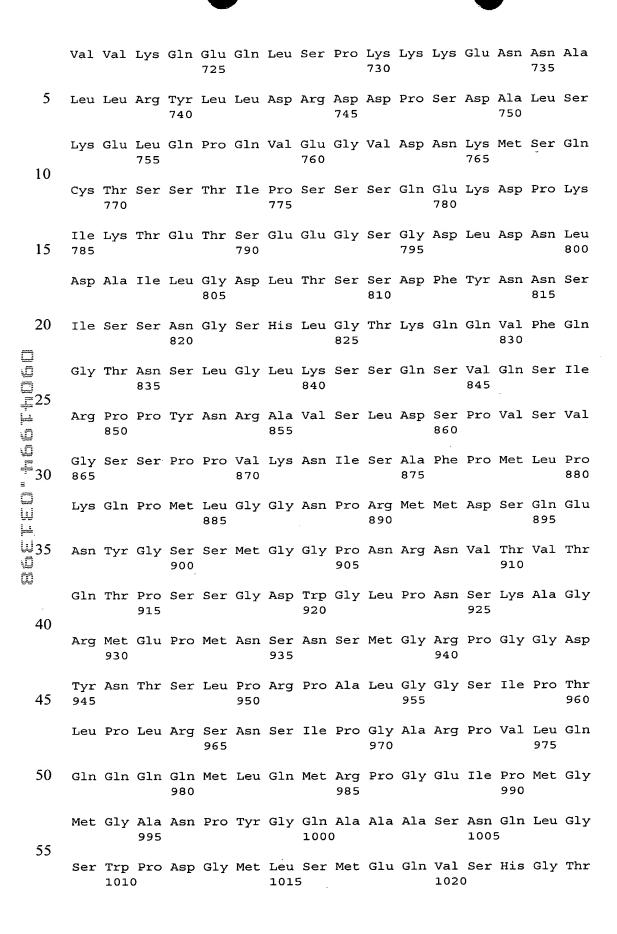
5			ATG Met														2800
3			AAT Asn														2848
10			AAC Asn														2896
15	Met	Gly	AGA Arg 940	Pro	Gly	Gly	Asp	Tyr 945	Asn	Thr	Ser	Leu	Pro 950	Arg	Pro	Ala	2944
20	Leu	Gly 955	GGC Gly	Ser	Ile	Pro	Thr 960	Leu	Pro	Leu	Arg	Ser 965	Asn	Ser	Ile	Pro	2992
고 고 고 주 25			AGA Arg														3040
e T			GAA Glu													Ala	3088
_ 30 			TCT Ser		Gln					Pro					Ser		3136
₩ ₩35 ₩ ₩			GTT Val 1020	Ser					Asn					Arg			3184
40			GAT Asp					Pro					Gly				3232
45		Arg	GCA Ala				Gln					Leu					3280
			GGC Gly			Glu					Leu					Leu	3328
50			CAG Gln		Gln					Lys					Gln		3376
55			GCA Ala 1100	Ala					Gln					Tyr			3424

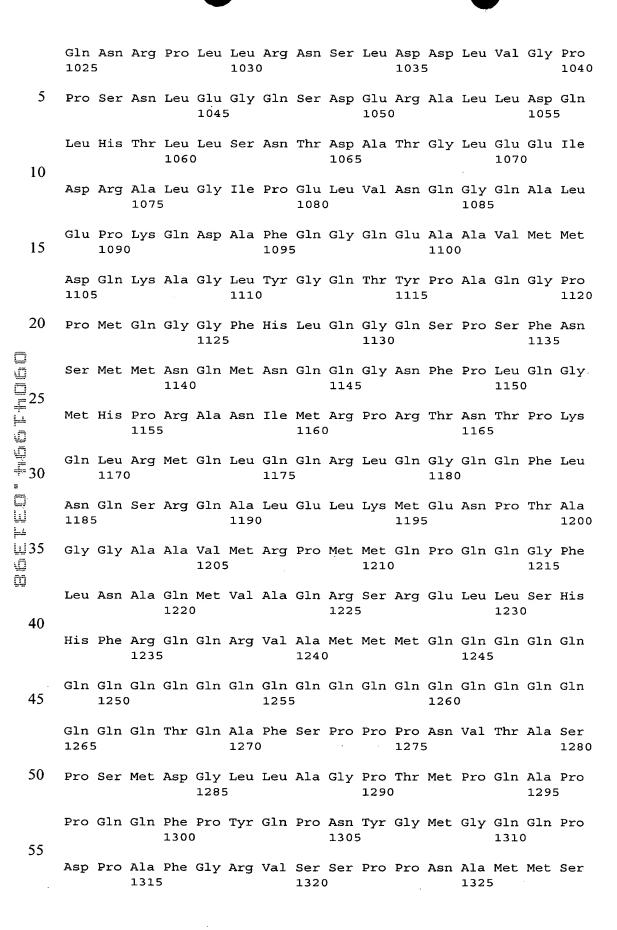
				Pro		CAG Gln			Pro					Phe				3472
	5		Gln			TCT Ser		Asn					Gln					3520
	10					CTC Leu 1150	Gln					Arg					Arg	3568
	15					ACC Thr					Arg					Gln		3616
	20				Gln	CAG Gln				Gln					Leu			3664
				Glu		CCT Pro			Gly					Met				3712
1000 - 10	25 = 25 = 30		Gln			CAG Gln		Phe					Met					 3760
## 10 #################################	30					CTA Leu 1230	Ser					Gln					Met	3808
	≟ 35					CAG Gln					Gln					Gln		3856
in the second se	40				Gln	CAG Gln				Gln					Phe			3904
				Asn		ACT Thr			Pro					Leu				3952
	45		Thr			CAA Gln		Pro					Pro					4000
	50					CAA Gln 1310	Gln					Phe					Ser	4048
	55					ATG Met					Met					Asn		4096
		ATG	ATG	CAA	CAC	CCG	CAG	GCT	GCA	TCC	ATC	TAT	CAG	TCC	TCA	GAA	ATG	4144

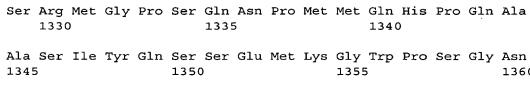
	Met Met Gln His Pro Gln Ala Ala Ser Ile Tyr Gln Ser Ser Glu Met 1340 1345 1350										
5	AAG GGC TGG CCA TCA GGA AAT TTG GCC AGG AAC AGC TCC TTT TCC CAG Lys Gly Trp Pro Ser Gly Asn Leu Ala Arg Asn Ser Ser Phe Ser Gln 1355 1360 1365	4192									
10	CAG CAG TTT GCC CAC CAG GGG AAT CCT GCA GTG TAT AGT ATG GTG CAC Gln Gln Phe Ala His Gln Gly Asn Pro Ala Val Tyr Ser Met Val His 1370 1375 1380 1385	4240									
15	ATG AAT GGC AGC AGT GGT CAC ATG GGA CAG ATG AAC ATG AAC CCC ATG Met Asn Gly Ser Ser Gly His Met Gly Gln Met Asn Met Asn Pro Met 1390 1395 1400	4288									
	CCC ATG TCT GGC ATG CCT ATG GGT CCT GAT CAG AAA TAC TGC TGA CAT CT Pro Met Ser Gly Met Pro Met Gly Pro Asp Gln Lys Tyr Cys * 1405 1410 1415	4338									
20	CTGCACCAGG ACCTCTTAAG GAAACCACTG TACAAATGAC ACTGCACTAG GATTATTGGG	4398									
	AAGGAATCAT TGTTCCAGGC ATCCATCTTG GAAGAAAGGA CCAGCTTTGA GCTCCATCAA	4458									
口 口 二 二 二 二 二 二 二 二 二 二 二 二 二 二 二 二 二 二	GGGTATTTTA AGTGATGTCA TTTGAGCAGG AATTCTAG	4496									
	(2) INFORMATION FOR SEQ ID NO:2:										
#30 U U U U U U U U U U	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 1417 amino acids(B) TYPE: amino acid(D) TOPOLOGY: linear										
<u> </u>	(ii) MOLECULE TYPE: protein										
1	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:										
40	Met Ser Gly Leu Gly Glu Asn Leu Asp Pro Leu Ala Ser Asp Ser Arg 1 5 10 15										
	Lys Arg Lys Leu Pro Cys Asp Thr Pro Gly Gln Gly Leu Thr Cys Ser 20 25 30										
45	Gly Glu Lys Arg Arg Glu Gln Glu Ser Lys Tyr Ile Glu Glu Leu 35 40 45										
	Ala Glu Leu Ile Ser Ala Asn Leu Ser Asp Ile Asp Asn Phe Asn Val 50 55 60										
50	Lys Pro Asp Lys Cys Ala Ile Leu Lys Glu Thr Val Arg Gln Ile Arg 65 70 75 80										
55	Gln Ile Lys Glu Gln Gly Lys Thr Ile Ser Asn Asp Asp Asp Val Gln 85 90 95										



Tyr Gly Leu Ala Asp Pro Ser Thr Thr Gly Gln Met Ser Gly Ala Arg 425 Tyr Gly Gly Ser Ser Asn Ile Ala Ser Leu Thr Pro Gly Pro Gly Met Gln Ser Pro Ser Ser Tyr Gln Asn Asn Asn Tyr Gly Leu Asn Met Ser 10 Ser Pro Pro His Gly Ser Pro Gly Leu Ala Pro Asn Gln Gln Asn Ile 475 Met Ile Ser Pro Arg Asn Arg Gly Ser Pro Lys Ile Ala Ser His Gln 15 485 490 Phe Ser Pro Val Ala Gly Val His Ser Pro Met Ala Ser Ser Gly Asn 505 Thr Gly Asn His Ser Phe Ser Ser Ser Leu Ser Ala Leu Gln Ala 515 520 525 Ile Ser Glu Gly Val Gly Thr Ser Leu Leu Ser Thr Leu Ser Ser Pro 535 **=**25 ᆂ Gly Pro Lys Leu Asp Asn Ser Pro Asn Met Asn Ile Thr Gln Pro Ser . 0 550 555 Lys Val Ser Asn Gln Asp Ser Lys Ser Pro Leu Gly Phe Tyr Cys Asp 565 Gln Asn Pro Val Glu Ser Ser Met Cys Gln Ser Asn Ser Arg Asp His Ų 580 585 Leu Ser Asp Lys Glu Ser Lys Glu Ser Ser Val Glu Gly Ala Glu Asn Q 595 600 605 Ø Gln Arg Gly Pro Leu Glu Ser Lys Gly His Lys Lys Leu Leu Gln Leu 40 Leu Thr Cys Ser Ser Asp Asp Arg Gly His Ser Ser Leu Thr Asn Ser Pro Leu Asp Ser Ser Cys Lys Glu Ser Ser Val Ser Val Thr Ser Pro 45 645 Ser Gly Val Ser Ser Ser Thr Ser Gly Gly Val Ser Ser Thr Ser Asn 665 50 Met His Gly Ser Leu Leu Gln Glu Lys His Arg Ile Leu His Lys Leu 675 680 Leu Gln Asn Gly Asn Ser Pro Ala Glu Val Ala Lys Ile Thr Ala Gln 690 55 Ala Thr Gly Lys Asp Thr Ser Ser Ile Thr Ser Cys Gly Asp Gly Asn 710 715







Leu Ala Arg Asn Ser Ser Phe Ser Gln Gln Gln Phe Ala His Gln Gly 1365 1370 1375

10 Asn Pro Ala Val Tyr Ser Met Val His Met Asn Gly Ser Ser Gly His 1380 1385 1390

Met Gly Gln Met Asn Met Asn Pro Met Pro Met Ser Gly Met Pro Met 1395 1400 1405

Gly Pro Asp Gln Lys Tyr Cys * 1410 1415